A New Chemometric Strategy Based on ^{1}H-NMR Data Applied for Authentication of Romanian Vegetable Oils

MIHAELA MIHALACHE[1], AURELIA BRATU[1]*, ANAMARIA HANGANU[2], NICOLETA-AURELIA CHIRA[1], MARIA-CRISTINA TODASCA[1], SORIN ROSCA[1]

[1] University Politehnica of Bucharest, Faculty of Applied Chemistry and Materials Science, "Costin D. Nenitzescu" Organic Chemistry Department, 1-7 Polizu Str., 011061, Bucharest, Romania

[2] Romanian Academy, Organic Chemistry Center "Costin D. Nenitzescu" 202B Splaiul Independentei, 060023, Bucharest, Romania

^{1}H-NMR spectroscopy coupled with Principal Component Analysis (PCA) as chemometrical method was used to authenticate Romanian vegetable oils (sunflower, soybean, linseed and rapeseed) according to their variety. ^{1}H-NMR spectroscopy offers two possibilities for vegetable oils authentication. The first one, based on ^{1}H-NMR integrals and systems of chemometric equations allowed to determine the composition of vegetable oil on four classes of fatty acids, which was further used as input for the PCA analysis in order to separate the oils according to their variety. The second possibility in terms of authentication is to transform the ^{1}H-NMR spectra into series of numerical values (vectors) by integrating certain peaks or spectral windows. This strategy, investigated in the present paper, showed good results in discriminating among sunflower, soybean, rapeseed and linseed oils.

Keywords: vegetable oils, ^{1}H-NMR, PCA, authentication

The search for the origin and the authenticity of food products has been the object of numerous studies in the past few years using various physico-chemical determinations which provide data that are further processed by chemometric methods [1-5].

In recent years the Romanian production of vegetable oils (sunflower, linseed, rapeseed and soybean) has increased significantly. Because of their chemical composition and nutritional values, vegetable seeds are considered to be a great source of lipids and proteins [6] and are largely used in the production of edible oils [7] and animal feed [8].

^{1}H-NMR spectroscopy is a powerful analytical tool for analysis and characterization of food products. In previous studies of our research group, ^{1}H-NMR spectroscopy data were used for identification and quantitative measurement of compounds usually found in vegetable oils [3, 9] and wines [10-13].

In this study we aimed to authenticate vegetable oils (sunflower, linseed, rapeseed and soybean) according to their variety using ^{1}H-NMR spectroscopy coupled with PCA method *(Principal Component Analysis)*.

^{1}H-NMR spectroscopy offers two ways in which data can be used to authenticate vegetable oils. The first way was to process by PCA the composition of vegetable oils on four classes of fatty acids (tri-unsaturated, di-unsaturated, mono-unsaturated and saturated fatty acids) obtained based on ^{1}H-NMR data and a system of chemometric equations [14]. The second strategy we investigate in this work consists in transforming the ^{1}H-NMR spectra into series of numerical values taking into account directly the integral value of selected peaks.

Experimental part

The vegetable oils were obtained in our laboratory by Soxhlet extraction [15] from certified oil seeds (in terms of variety, geographical origin, crop year). The seed samples were provided by the Academy of Agricultural Sciences of Bucharest, the National Institute for Agricultural Research and Development of Fundulea, and regional Research Stations for Agricultural Development. All oilseeds we used in our study were produced in Romania.

The oil sample was dissolved in $CDCl_3$ (2:8 v/v). The sample was sonicated for 5 minutes, for degassing and mixing.

^{1}H-NMR spectra were recorded on a Bruker Avance DRX 400 spectrometer, operating at 9.4 Tesla, corresponding to the resonance frequency of 400.13 MHz for the ^{1}H nucleus, equipped with a direct detection four nuclei probe head and field gradients on *z* axis. Samples were analyzed in 5 mm NMR tubes (Wilmad 507). The chemical shifts are reported in ppm, using the TMS as internal standard. Typical parameters for ^{1}H-NMR spectra were: 30° pulse, 4 s acquisition times, 6.4 KHz spectral window, 8 scans, 52 K data points. The FID was not processed prior to Fourier transformation. The average acquisition time of the ^{1}H-NMR spectra was approximately 2 min. The statistic analyses (*Principal Component Analysis*) used to investigate the compositional differences between oils (sunflower, soybean, linseed and rapeseed) was carried out using the *XLSTAT* software.

Results and discussions

A total amount of 115 samples of vegetable oils has been subjected to this study (17 samples of sunflower, 48 samples of soybean, 21 samples of linseed and 29 samples of rapeseed oil).

^{1}H-NMR spectra of vegetable oils have the same shape, but the differences between them are reflected on the integral values of the characteristic peaks. Figure 1 presents the ^{1}H-NMR spectrum of rapeseed oil and table 1 shows the chemical shifts and peak assignment of ^{1}H-NMR spectra of vegetable oils, according to the literature specifications [16].

Based on ^{1}H-NMR spectra and using systems of chemometric equations the composition of vegetable oils

* *email: rosuaurelia36@yahoo.com; Tel: 0722413051*

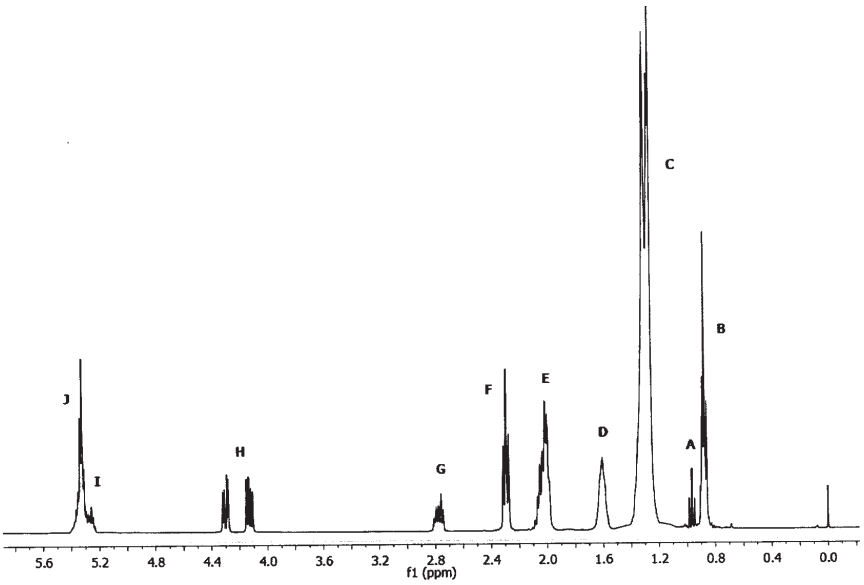


Fig.1. ^{1}H-NMR spectrum of rapeseed oil

Signal	δ (ppm)	Proton	Compound
A	0,95	-CH=CH-CH$_2$-C**H$_3$**	linolenic acid
B	0,85	-CH$_2$-CH$_2$-CH$_2$-C**H$_3$**	all acids except linolenic acid
C	1,25	-(C**H$_2$**)n-	all fatty acids
D	1,6	-C**H$_2$**-CH$_2$-OCO-	all fatty acids, β-methylene protons
E	2,02	-C**H$_2$**-CH=CH-	all unsaturated fatty acids
F	2,3	-C**H$_2$**-OCO-	all fatty acids, α-methylene protons
G	2,75	-CH=CH-C**H$_2$**-CH=CH-	linoleic acid and linolenic acid
H	4.19	-C**H$_2$**-OCOR	all fatty acids
I	5.24	-C**H**-OCOR	all unsaturated fatty acids
J	5.35	-C**H**=C**H**-	all unsaturated fatty acids

Table 1
CHEMICAL SHIFTS AND PEAK ASSIGNMENT OF ^{1}H-NMR SPECTRA OF VEGETABLE OILS

was determined in our previous studies [3,14] on four classes of fatty acids: tri-unsaturated, di-unsaturated, mono-unsaturated and saturated. The obtained composition was processed by mean of PCA method with good results regarding the authentication: a very good separation of the rapeseed and linseed oils was obtained, while the groups of sunflower and soybean oils were relatively close; therefore, another representation of principal component scores was necessary using only the data of sunflower and soybean oils in order to obtain a good discrimination of sunflower and soybean oils.

The current paper aims to ease the authentication process using the information directly provided by the ^{1}H-NMR spectra. Therefore each signal from the spectra was integrated in triplicate, for a good precision and the mean integral value was used in further computation. Ten values for each spectrum were obtained corresponding to the *A* to *J* signals of the ^{1}H-NMR spectra.

Table 2 presents the mean integral values obtained for each signal from the ^{1}H-NMR spectra for ten samples of each type of the analyzed vegetable oil.

Table 2
THE MEAN INTEGRAL VALUES OBTAINED FOR EACH SIGNAL FROM THE ^{1}H-NMR SPECTRA

Sample	The mean integral values									
	B	A	C	D	E	F	G	H	I	J
SF-1	3.00	0.00	15.57	2.53	1.39	1.94	0.21	1.27	0.38	0.86
SF-2	3.00	0.00	15.71	2.23	2.05	1.88	0.47	1.22	0.33	1.46
SF-3	3.00	0.00	16.60	1.98	3.39	1.93	1.23	1.24	0.92	2.44
SF-4	3.00	0.00	17.06	1.98	3.36	1.93	1.01	1.24	0.87	2.22
SF-5	3.00	0.00	17.00	1.98	3.44	1.93	1.10	1.22	0.89	2.32
SF-6	3.00	0.00	16.94	1.90	3.13	1.82	0.88	1.16	0.48	2.43
SF-7	3.00	0.00	16.45	1.97	3.39	1.91	1.14	1.22	0.73	2.46
SF-8	3.00	0.00	16.78	2.02	3.11	1.88	0.99	1.19	0.59	2.36
SF-9	3.00	0.00	16.38	1.97	3.22	1.85	1.07	1.20	0.65	2.43
SF-10	3.00	0.00	16.24	2.41	1.61	1.96	0.18	1.32	0.34	0.93

LIN-1	3.00	3.86	29.83	4.55	8.10	4.44	5.62	2.86	2.52	8.01
LIN-2	3.00	3.09	27.53	4.01	7.03	3.94	4.55	2.55	2.14	6.71
LIN-3	3.00	3.79	28.80	4.44	7.95	4.35	5.54	2.75	2.60	7.84
LIN-4	3.00	3.09	27.14	4.00	7.10	3.91	4.67	2.53	2.31	6.70
LIN-5	3.00	2.93	27.05	3.92	6.99	3.79	4.29	2.45	2.18	6.41
LIN-6	3.00	2.97	27.08	3.95	7.07	3.83	4.34	2.45	2.13	6.55
LIN-7	3.00	3.43	27.95	4.22	7.51	4.11	5.07	2.62	2.47	7.20
LIN-8	3.00	4.36	30.83	4.91	8.76	4.76	6.31	3.07	2.86	8.88
LIN-9	3.00	4.14	30.23	4.70	8.34	4.59	6.00	2.96	2.79	8.30
LIN-10	3.00	3.40	28.64	4.24	7.40	4.12	4.99	2.64	2.43	7.17
S-1	3.00	0.24	18.06	2.09	3.46	2.06	1.39	1.35	1.00	2.58
S-2	3.00	0.21	18.28	2.09	3.40	2.04	1.25	1.31	0.87	2.53
S-3	3.00	0.23	18.09	2.10	3.48	2.06	1.36	1.34	1.01	2.57
S-4	3.00	0.21	18.18	2.13	3.53	2.08	1.36	1.34	0.94	2.61
S-5	3.00	0.21	18.37	2.10	3.36	2.06	1.23	1.34	0.97	2.37
S-6	3.00	0.24	17.96	2.12	3.49	2.07	1.38	1.33	0.80	2.74
S-7	3.00	0.22	18.20	2.10	3.39	2.05	1.31	1.33	0.88	2.58
S-8	3.00	0.15	18.24	2.10	3.37	2.06	1.20	1.33	0.98	2.33
S-9	3.00	0.20	18.41	2.12	3.44	2.07	1.26	1.35	0.96	2.48
S-10	3.00	0.22	18.08	2.10	3.45	2.06	1.33	1.34	0.98	2.54
R-1	3.00	0.26	19.86	2.11	3.78	2.07	0.76	1.31	0.73	2.35
R-2	3.00	0.22	19.80	2.12	3.85	2.05	0.68	1.31	0.72	2.30
R-3	3.00	0.25	19.92	2.11	3.82	2.08	0.71	1.32	0.70	2.39
R-4	3.00	0.24	19.81	2.11	3.85	2.07	0.67	1.32	0.69	2.32
R-5	3.00	0.27	19.67	2.10	3.87	2.07	0.69	1.33	0.73	2.29
R-6	3.00	0.21	19.27	2.04	3.72	2.02	0.64	1.30	0.68	2.24
R-7	3.00	0.26	20.04	2.13	3.89	2.07	0.68	1.27	0.71	2.31
R-8	3.00	0.25	20.02	2.15	3.90	2.12	0.69	1.32	0.64	2.43
R-9	3.00	0.29	20.63	2.21	3.99	2.13	0.69	1.30	0.72	2.42
R-10	3.00	0.31	20.07	2.13	3.89	2.07	0.68	1.29	0.75	2.32

* SF - sunflower; LIN - linseed; R - repeseed; S - soybean

As it can be seen from table 2, the signal B has the same constant value (3.00) for all the samples because it was used as a calibration signal and therefore this signal is not used in further statistical computations.

The mean integral values of the *A, C, D, E, F, G, H, I* and *J* signals obtained for each 115 samples of the analyzed vegetable oils (17 samples of sunflower, 48 samples of soybean, 21 samples of linseed and 29 samples of rapeseed) were processed using PCA as statistical method.

Figure 2 shows the representation of principal component scores PC1/PC2 for the sunflower, linseed, soybean and rapeseed oils using ^{1}H-NMR integral values of *A, C, D, E, F, G, H, I* and *J* signals.

In figure 2 a very good discrimination of linseed oils from the other vegetable oils is observed using ^{1}H-NMR integral values of *A, C, D, E, F, G, H, I* and *J* signals. As it can be noticed, sunflower, soybean and rapeseed oils show a tendency to group in the same place; therefore, another representation was made excluding the linseed oils.

Figure 3 show the representation of principal component scores PC1/PC2 using the same integral values of *A, C, D, E, F, G, H, I* and *J* signals, for the sunflower, soybean and rapeseed oils excluding the linseed oils.

A very good discrimination of sunflower oil, rapeseed oil and soybean oil is observed from figure 3, but in this case sunflower oils group shows a large spreading area.

A better strategy of discrimination was designed in further calculations, by properly combining the integral values of a reduced number of peaks (*A, C, E, F, G, H* and *I*).

The integrals selection was made based on the calculation of the components relevance.

The integrals relevance is obtained by following several steps:

a) *calculating the average of the integrals for each oil variety* – a mean integral value is obtain for each signal in the ^{1}H-NMR spectrum

e.g. for linseed oil:

mean $A_{LIN} = (A_{LIN-1}+A_{LIN-2}+\ldots+A_{LIN-21})/21$

were:

A_{LIN-1} represents the integral value of signal *A*, in LIN-1 oil sample.

The same calculation is made for *B, C, D, E, F, G, H, I* and *J* integrals.

b) *calculating the integral variability within oil variety* – was obtained by the difference between the mean integral value (obtained at point (a)) and the value of the integral for each oil sample and represents the variability of the oil sample integral from the mean value.

$\text{var } A_{LIN-1} = \text{abs}(\text{mean } A_{LIN} - A_{LIN-1})$

$\text{mean var } A_{LIN} = (\text{var } A_{LIN-1} + \text{var } A_{LIN-2} + \ldots + \text{var } A_{LIN-21})/21$

were:

$\text{var } A_{LIN-1}$ represents the variability of the *A* integral of LIN-1 oil sample within the lin oil variety;

$\text{mean var } A_{LIN}$ represents the integral variability of linseed oil within the oil variety.

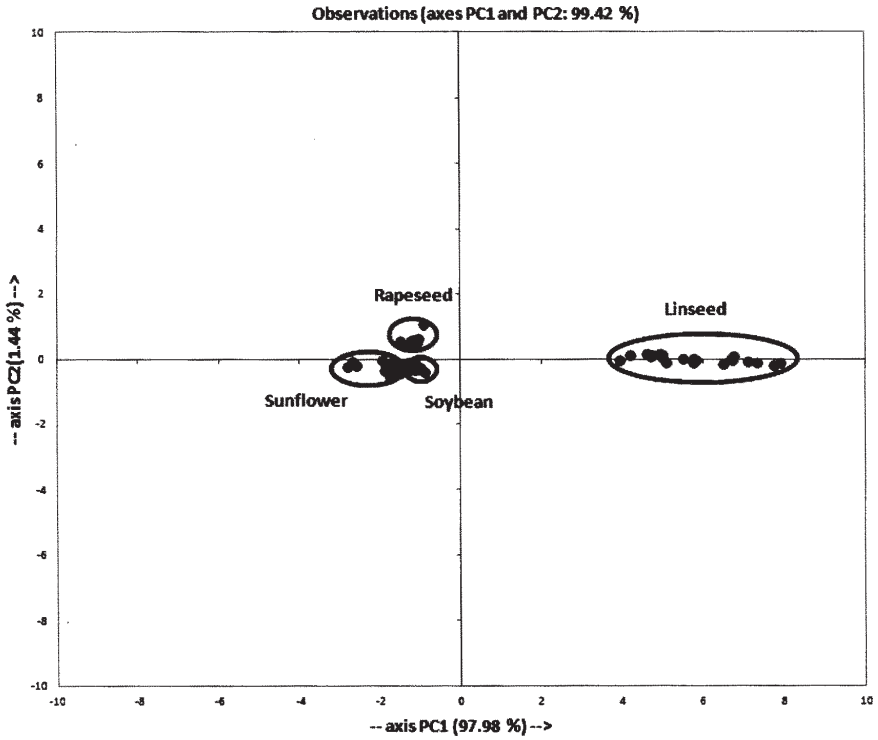


Fig.2. Principal Component scores PC1/PC2 plot for the sunflower oil, rapeseed oil, soybean oil and linseed oil using ^{1}H-NMR integral values of *A, C, D, E, F, G, H, I* and *J* signals statistically processed by PCA

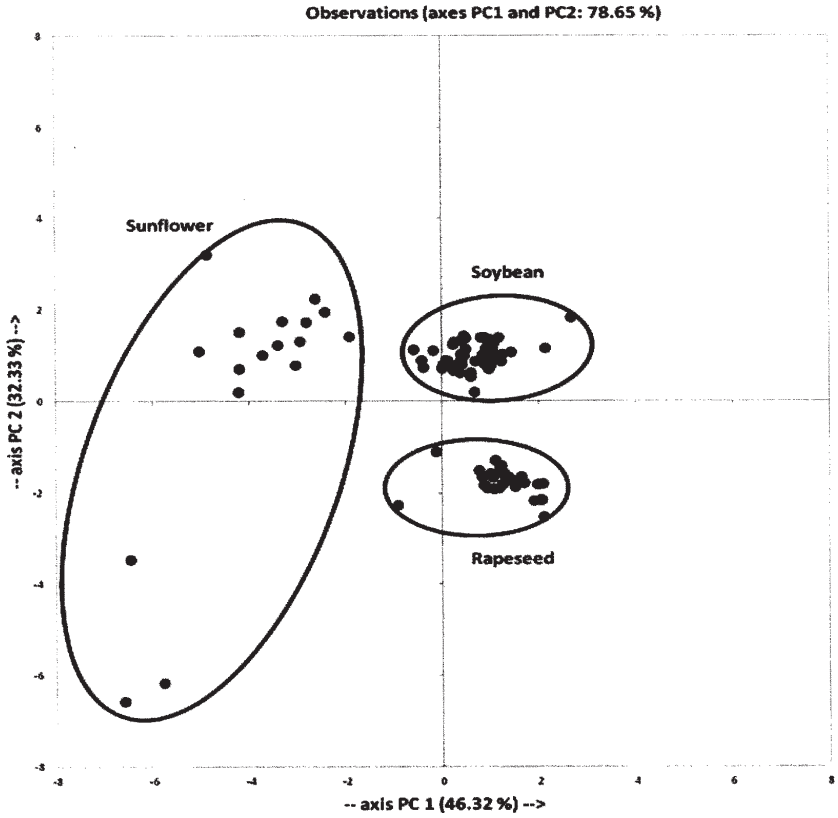


Fig.3. Principal Component scores PC1/PC2 plot for the sunflower oil, rapeseed oil and soybean oil using ^{1}H-NMR integral values of *A, C, D, E, F, G, H, I* and *J* signals statistically processed by PCA

c) *calculating the average of each mean integral value between oil varieties*

$M_A = (\text{mean } A_{LIN} + \text{mean } A_{SUNFLOWER} + \text{mean } A_{SOY} + \text{mean } A_{RAPE})/4$

were:

M_A = mean *A* integral value between oil varieties.

The same calculation is made for *B, C, D, E, F, G, H, I* and *J* integrals.

d) *calculating the integral variability between the oil varieties* – was obtained by difference between the average of each mean integral value (M) and the average of integrals for each oil variety:

$\text{Variability } A_{LIN} = \text{abs}(M_A - \text{mean } A_{LIN})$

e) *calculating the integrals relevance for each oil variety (Rel)*

$\text{Rel } A_{LIN} = (\text{Variability } A_{LIN})/(\text{mean var } A_{LIN})$

f) *calculating the total relevance for each integral (RT)*

$RT_A = \text{Rel } A_{LIN} + \text{Rel } A_{SUNFLOWER} + \text{Rel } A_{SOY} + \text{Rel } A_{RAPE}$

Based on the model described, the total relevance for all the integrals was calculated. Table 3 presents the total relevance for all the integrals, in ascending order.

As it can be noticed from table 3, integral *B* is not a relevant component, followed by *D* and *J*, therefore is

Table 3
INTEGRALS TOTAL RELEVANCE IN ASCENDING ORDER

Integrals	Total relevance
B	0.00
D	2.50
J	2.76
H	5.40
A	7.41
I	8.39
F	8.62
E	9.45
C	10.28
G	15.09

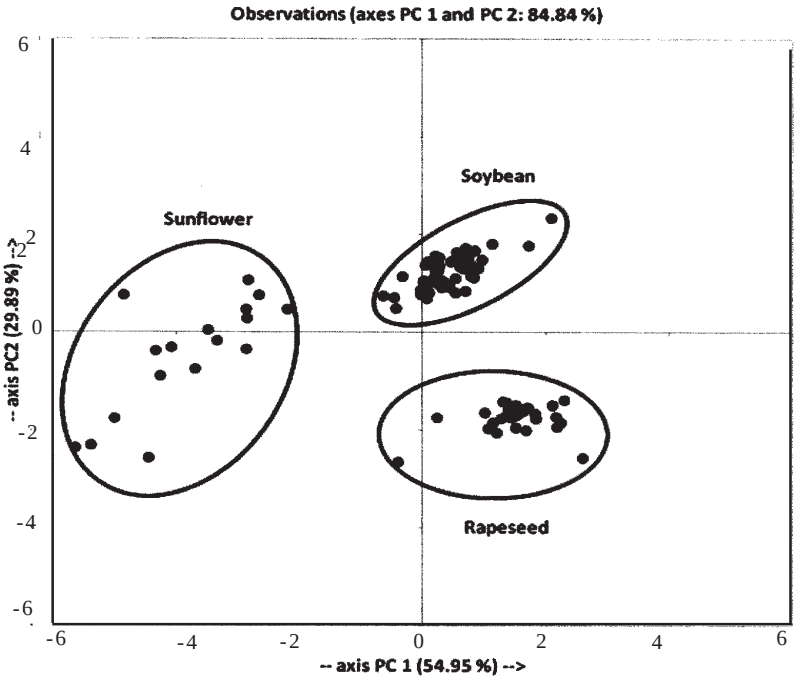


Fig.4. Principal Component scores PC1/PC2 plot for the sunflower oil, rapeseed oil and soybean oil using ^{1}H-NMR integral values of *A, C, E, F, G, H* and *I* signals statistically processed by PCA

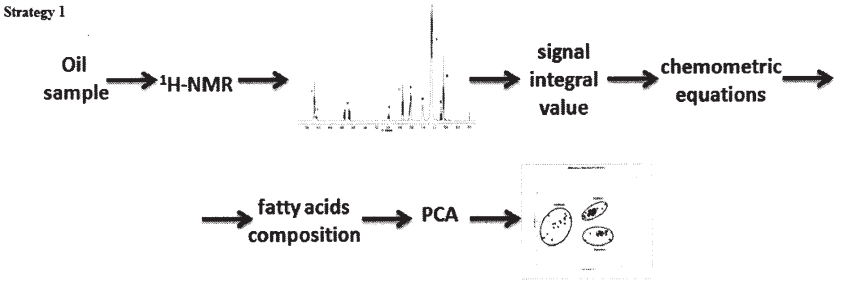
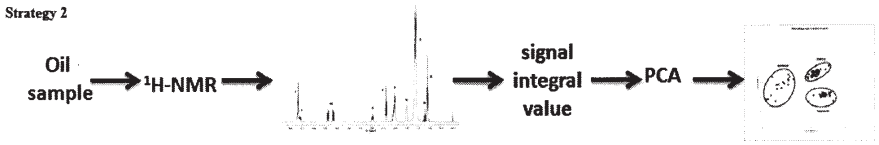


Fig. 5. Authentication strategies



proceed to the gradual elimination of components with low relevance.

Figure 4 presents an improvement of the grouping tendency by plotting of principal component scores PC1/PC2 for the sunflower, soybean and rapeseed oils using ^{1}H-NMR integral values of *A, C, E, F, G, H* and *I* signals.

As it can be seen from figure 4 an improvement of the grouping tendency is observed in this case.

This strategy helps us to obtain a very good authentication of the vegetable oils using the integral for only 7 selected peaks.

We choose to eliminate peak B (corresponding to protons from all methyl groups from all acids except linolenic acid), peak D (corresponding to β-methylene protons with respect to the carbonyl group from all acids) and peak J (corresponding to protons linked with unsaturated carbons from all unsaturated fatty acids). These 3 variables made the authentication approach to be considered from the beginning an unproductive way.

A scheme containing the proposed two strategies procedures for studying the authentication of vegetable oils is presented in figure 5.

Conclusions

^{1}H-NMR spectroscopy coupled with *Principal Component Analysis* chemometrical method was used to authenticate vegetable oils. The advantage of ^{1}H-NMR spectroscopy is, besides being fast, that it can be applied directly on triglycerides, without any sample preparation.

By using both authentication strategies presented in figure 5, satisfactory results were obtained.

The second strategy applied to the vegetable oils is faster compared to the first one and gives better results in terms of authentication but it does not provide compositional results.

The obtained conclusions are important for selecting the proper method for authentication studies.

Acknowledgement: The work has been funded by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Romanian Ministry of Labour, Family and Social Protection through the Financial Agreement POSDRU/88/1.5/S/60203.

References

1. GALTIER, O., DUPUY, N., DREAU, Y.L., OLLIVIER, D., PINATEL, C., KISTER, J., ARTAUD, J., Anal. Chim. Acta, **595**, 2007, p.136.
2. MAGGIO, R.M., CERRETANI, L., CHIAVARO, E., KAUFMAN, T.S., BENDINI, A., Food Control, **21**, 2010, p.890.
3. CHIRA, N.A., TODASCĂ, M.C., PAUNESCU, G., DAVID, I., IONESCU, N. STANCIU, M., ROSCA, S., Rev. Chim.(Bucharest), **62**, no. 4, 2011, p.391.
4. GURDENIZ, G., OZEN, B., Food Chem., **116**, 2009, p.519.
5. SMEJKALOVA, D., PICCOLO, A., Food Chem., **118**, 2010, p.153.
6. OMIDI, H., TAHMASEBI, Z., BADI, H.A.N., TORABI, H., MIRANSARI, M., C.R. Biol., **333**, 2010, p.248.
7. NIMETA, G., DA SILVA, E.A., PALUA, F., DARIVAC, C., DOS SANTOS FREITASC, L., NETOB, A.M., FILHOB, L.C., Chem. Eng. J., **168**, 2011, p.262.
8. DEREAU, M., AUROUSSEAU, E., MARTIN, C., Anim. Feed Sci. Teh., 150, 2009, p.187.
9. HANGANU, A., TODASCĂ, M.C., CHIRA, N.A., MAGANU, M., ROSCA, S., Food Chem., **134**, 2012, p.2453.
10. HANGANU, A., TODASCĂ, M.C., CHIRA, N.A., ROSCA, S., Rev. Chim. (Bucharest), **62**, no. 7, 2011, p.689.
11. HANGANU, A., TODASCĂ, M.C., CHIRA, N.A., ROSCA S., U.P.B Sci. Bull., **73 (2)**, 2011, p.197.
12. TODASCĂ, M.C., FOTESCU, L., CHIRA, N.A., DELEANU, C., ROSCA, S., Rev. Chim.(Bucharest), **62**, no**.** 2, 2011, p.131.
13. TODASCĂ, M.C., FOTESCU, L., HÎNCU, F.A., HANGANU, A., CHIRA, N.A., ROSCA, S. Rev. Chim. (Bucharest), **61**, no. 11, 2010, p.1042.
14. CHIRA, N.A., TODASCĂ, M.C., NICOLESCU, A., ROSU, A., NICOLAE, M., ROSCA, S.I., Rev. Chim.(Bucharest), **62**, no. 1, 2011, p.42.
15. *** EN ISO 659: 2003, Oilseeds. Determination of Oil Content.
16. CHIRA, N.A., TODASCĂ, M.C., NICOLESCU, A., PĂUNESCU, G., ROSCA, S.I., U.P.B Sci. Bull., **71 (4)**, 2009, p.3

Manuscript received: 2.05.2012